

COVER PAGE

Hewlett-Packard Company Docket Number:

100110642-1

Title:

Centralized Scalable Resource Architecture and System

Inventors:

Byron A. Alcorn
2731 Maroon Ct.
Fort Collins, Colorado 80525

Kevin T. Lefebvre
2707 Whitworth Dr.
Fort Collins, Colorado 80525

CENTRALIZED SCALABLE RESOURCE ARCHITECTURE AND SYSTEM

TECHNICAL FIELD OF THE INVENTION

[0001] The present invention relates generally to the field of computer systems, and more particularly to a centralized scalable resource architecture and system.

BACKGROUND OF THE INVENTION

[0002] In the days when mainframe computers and super computers dominated the digital landscape, users were typically given slices of computing time from a fixed pool so that the computer worked on many jobs in parallel. Users interfaced with the mainframe computers using "dumb terminals" connected to the mainframe computers and a simple command-line user interface. However, the computing environment gradually shifted to lower-cost personal computers and workstations, which offers users greater control of how their applications are executed and a better and more intuitive graphical user interface. More recently however, with the global connectivity provided by the Internet and faster bandwidths, there has been another shift toward simplified or stripped-down computers or appliances connected with the centralized data storage and computing resources provided by "web farms" via the Internet. Web farms are typically clusters of servers, microcomputers and mainframe computers that currently perform mostly web server and hosting function for web applications and web pages.

[0003] In the computer graphics domain, current computer graphical visualization systems may employ a cluster of workstations or personal computers executing a graphics application and working simultaneously on a single job to render an image to be displayed on a display device. The image is displayed as a single logical image on a single monitor screen or across multiple monitor screens. Such

visualization systems may use a plurality of graphics pipelines to render different portions of an image for display on the display monitor to speed up processing time and improve the quality of the displayed image.

SUMMARY OF THE INVENTION

[0004] In accordance with an embodiment of the present invention, a centralized resource system comprises a plurality of compute resource units, a plurality of visualization resource units, and a switching fabric coupling the plurality of visualization resource units to the plurality of compute resource units. The switching fabric is operable to couple select one or more visualization resource units to select one or more compute resource units for generating one or more graphical images. A plurality of display devices are operable to display the one or more graphical images coupled to the one or more select visualization resource units.

[0005] In accordance with another embodiment of the invention, a centralized resource system comprises a plurality of first compute resource units, a plurality of second compute resource units, and a switching fabric coupling the plurality of second compute resource units to the plurality of first compute resource units. The switching fabric is operable to selectively couple outputs of the plurality of first compute resource units to inputs of the plurality of second compute resource units. The plurality of first and second compute resource units operable and function together to generate one or more execution results. A plurality of display devices is coupled to the plurality of first and second compute resource units and operable to receive execution results therefrom.

[0006] In accordance with yet another embodiment of the present invention, a centralized resource system comprises first resource means, second resource means, and means for selectively coupling one or more outputs of the first resource means to one or more inputs of the second resource means. A plurality of display means is coupled to the first and second resource means and operable to receive and display execution results therefrom.

[0007] In accordance with another embodiment of the present invention, a method comprises the steps of receiving a graphics visualization job to be executed by a plurality of compute resources, determining compute resource

requirements for the job, determining compute resource availability, and allocating compute resources in response to the determined compute resource requirements and availability. The method further comprises determining destinations to receive results of the job, and allocating and configuring communication channels from the allocated compute resources to the destinations.

[0008] In accordance with a further embodiment of the present invention, a graphics visualization architecture comprises a plurality of compute resource units, a plurality of graphics pipelines, and a first switching fabric coupling the plurality of graphics pipelines to the plurality of compute resource units. The first switching fabric is operable to selectively couple outputs of the plurality of compute resource units to inputs of the plurality of graphics pipelines. The architecture further comprises a plurality of compositors, and a second switching fabric coupling the plurality of compositors to the plurality of graphics pipelines, where the second switching fabric is operable to selectively couple outputs of the plurality of graphics pipelines to inputs of the plurality of compositors.

BRIEF DESCRIPTION OF THE DRAWINGS

[0009] For a more complete understanding of the present invention, the objects and advantages thereof, reference is now made to the following descriptions taken in connection with the accompanying drawings in which:

[0010] FIGURE 1 is a simplified diagram of an embodiment of a centralized scalable resource architecture and system according to the teachings of the present invention;

[0011] FIGURE 2 is a simplified block diagram of an embodiment of a centralized scalable resource architecture and system for graphics visualization according to the teachings of the present invention;

[0012] FIGURE 3 is a more detailed block diagram of an embodiment of a centralized scalable resource architecture and system for graphics visualization according to the teachings of the present invention;

[0013] FIGURE 4 is a simplified block diagram of a graphics engine; and

[0014] FIGURE 5 is a flowchart of a process for determining resource allocation according to an embodiment of the present invention.

DETAILED DESCRIPTION OF THE DRAWINGS

[0015] The preferred embodiment of the present invention and its advantages are best understood by referring to FIGURES 1 through 5 of the drawings, like numerals being used for like and corresponding parts of the various drawings.

[0016] Current computer graphical visualization systems such as Hewlett-Packard Company's Visualization Center sv6 (HP sv6) typically employ a cluster of workstations and visualization resources processing and rendering a single graphical image to be displayed as a single image. HP sv6 is described in U.S. Patent Application, Serial No. 09/715,335, entitled "SYSTEM AND METHOD FOR EFFICIENTLY RENDERING GRAPHICAL DATA," filed on November 17, 2000.

[0017] FIGURE 1 is a simplified diagram of an embodiment of a system 10 constructed according to the centralized scalable resource architecture of the present invention. System 10 employs a centralized resource 12, which may comprise data storage building blocks, computing resource building blocks, and other scalable resources. For example, centralized resource 12 comprises personal computers, workstations, servers and other computers 14 that can be easily added to the configuration of computing resources or removed therefrom according to need. A software such as a graphics application is executed in computers 14 acting as host computers. Centralized resource 12 may also comprise more specialized computing resources such as graphics or visualization resources 16 coupled to computers 14. Visualization resource 16 may comprise graphics pipelines, frame buffers, compositors and other specialized hardware and software resources used for graphics rendering. Visualization resource 16 may be implemented in hardware ranging from specialized graphics cards to powerful workstations. Centralized resources 12 are coupled to user terminals or display devices 18 via a network 20. Network 20 may be a local area network (LAN), wide area network (WAN), Internet, or any other suitable network or connectivity. User terminals or display devices 18 are devices on which users may view results data processed and generated by centralized resources 12. Display devices 18 may employ any display technology now known or later

developed. According to the teachings of the present invention, all the processing and visualization functions are performed by centralized resources 12 so that only data associated with the processed result are conveyed to the user. For example in the graphics application, only pixel data associated with the rendered 2-dimensional (2D) or 3-dimensional (3D) images to be displayed on display devices 18 are transmitted over network 20. The pixel or image data to be displayed to the user is usually several magnitudes less than the data needed to generate the graphics images, so that the data transmission time is significantly shortened and the data storage requirement at the users' computer terminals is minimal.

[0018] Unlike conventional systems, centralized scalable system 10 is operable to allocate centralized resources 12 to multiple jobs simultaneously to generate multiple images to be displayed on the respective display devices. The allocation unit of the computing resources, for example, may be a CPU (central processing unit) and the allocation unit of the visualization resources, for example, may be a graphics pipeline and its associated hardware/software. Referring to FIGURE 2, a subset of resources 22 is allocated to each job or an execution of an application software for displaying results on a subset of display devices 24, so that centralized scalable resource system 10 is operable to perform multiple jobs and execute multiple applications. Subset of resources 22 may comprise one or more networked data storage units 26, one or more computing resource units 28, and one or more visualization resource units 30. Centralized and scalable resource architecture and system 10 comprises a configurable connectivity between centralized resources 12 and display devices 18 to enable re-allocation on-the-fly.

[0019] FIGURE 3 is a more detailed block diagram of an embodiment of a centralized scalable resource architecture and system 10 for graphics visualization according to the teachings of the present invention. Centralized scalable resource architecture and system 10 comprises computing resources 14 such as a plurality of application hosts coupled to a switching fabric 34, which provides configurable connectivity to a graphics engine 36 of visualization resources 16. Not explicitly shown in this figure is a cluster of data storage devices that may be coupled to application hosts 14 for storing the data needed for generating the graphical images. Referring to FIGURE 4, graphics engine 36 comprises a plurality of workstations

with graphics pipelines 50 receiving graphical data 52 from application hosts 14, where switching fabric 34 determines which graphics pipelines receive graphical data from which application hosts. Graphics pipelines 50 may be arranged to have a master pipeline and slave pipelines interconnected by a local area network (LAN), for example. Depending on the resource requirements of an application and the size of the job, the number of application hosts and the number of graphics pipelines working on the job can be configured by changing the connectivity of the application hosts to the graphics pipelines. Switching fabric 34 may comprise switches, routers, switched networks and/or other network components. In particular, switching fabric 34 comprises a $M \times N$ switch that can couple any of its inputs to any of its outputs, such as a crossbar switch, a cross-connect, an Ethernet switch, or any device which is operable to connect any of its outputs to its inputs on-the-fly. The number of switch inputs, M , may equal the number of switch outputs, N . A frame buffer (FB) 54 is coupled to each graphics pipeline 50 and operable to store one or more frames of image or pixel data rendered by the graphics pipeline. The output of frame buffers 54 is preferably DVI (digital video interface) data 56 but may be other types of graphics data. DVI data typically comprise the coordinate values of the pixel position to be displayed on the display screen, and the color values of the pixel. For example, DVI data may comprise the (X,Y) coordinate value of the pixel, RGB (red, green, blue) values. DVI data may further comprise a depth value for the pixel (Z), and a transparency value for the pixel (α).

[0020] Referring also to FIGURE 3, each output from a frame buffer 54 is coupled to a converter 38. A second switching fabric 40 is coupled to the output of converters 38 to route the IP packets to the proper subset of compositors 42, IVA (Internet visualization architecture) components 44, and the display devices beyond for display to the users.

[0021] It should be noted that FIGURE 3 is a functional block diagram and that functionality of more than one block shown therein may be integrated together. For example, the functionality of IP Converter 38 can be integrated with graphics engine 36, or that the functionality of compositors 42 may be incorporated into IVA 44 or switching fabric 40.

[0022] In operation, application hosts 14 execute one or multiple graphics applications or one or more instances of a graphics application to render one or multiple graphical images. The graphical images may be displayed on one or more display devices. Switching fabric 34 is operable to connect, on-the-fly, the output of selected application hosts to selected graphics pipelines in graphics engine 36. Therefore, the cluster of servers, workstations and computers is essentially configurable into one or more sub-clusters to process one or more jobs. The output of the application host(s) executing a particular job can be coupled and routed to one or more graphics pipelines for rendering the image by switching fabric 34. IP converters 38, receiving the output from graphics engine 36, are operable to convert or packetize DVI data into data packets to facilitate transmission over a network. Depending on the type of network used to transmit the data to the display devices, converter 38 inserts header and protocol information according to the network type. For example, Transmission Control Protocol (TCP) and Internet Protocol (IP) may be used as the transmission protocol so that converters 38 are operable to convert the pixel data from the pipelines in graphics engine 36 to IP packets. Second switching fabric 40 is operable to connect and route, on-the-fly, the output of selected converters to the input of selected compositors 42. Each compositor 42 is operable to combine or composite one or more image data streams from converters 38 into one or more image data streams that are then sent to the display devices for display. Because the data are in the form of data packets or IP packets, for example, compositors 42 are operable to process the data after stripping off the header and other information. Thereafter, IVA 44 is operable to compress the composited image data, packetize it, and send it to one or more remote receivers or client devices for display via a network such as the Internet and other suitable networks. Therefore, IVA is operable to add destination information, protocol headers and other data to the image data to form data packets. IVA is also responsible for transmitting user input data such as keyboard and mouse input entered by the user to application host 14. Such user input may be used to control the processing of application hosts 14 and/or graphics engine 36. The users entering such input may be co-located with one or more of the display devices or be located remotely from the display devices.

[0023] In this manner, any number of compute resource units may operate together to process a single job, resulting in the possibility of together executing multiple jobs by centralized scalable resource system 10 to be sent to multiple destinations. The allocation of the compute resource units and the visualization resource units is determined based on a number of considerations. FIGURE 5 is a flowchart of a controlling process 60 for determining and making resource allocation according to an embodiment of the present invention. Resource allocation method 60 may employ agent technology. An agent may be software and/or hardware that accepts jobs submitted by the users, as shown in block 62. The agent determines where the data needed for executing the job is stored, as shown in block 64, which may be multiple storage devices and/or multiple storage locations. The agent then determines the compute resource requirements for the job and the availability of the needed compute resources, as shown in blocks 66 and 68. For graphical applications, the agent also determines the visualization resource requirements and the availability thereof, as shown in blocks 70 and 72. The agent then allocates the resources, as shown in block 74. The amount of resources allocated may be less than the optimum required for the submitted job due to resource unavailability. Next in block 76 the agent determines and allocates the communication channels needed to connect the data storage to the allocated compute resources and to the visualization resources. The agent also determines and allocates the communication channels to the users or the display devices. The agent configures the switching fabrics to selectively connect the appropriate allocated compute and visualization resources and compositors. The agent then initiates the job execution or sends a message to the application hosts to initiate job execution. The process ends in block 80.

[0024] Centralized scalable resource system and architecture 10 may be used for generalized computing applications or specialized applications such as graphics and visualization. The resources can be remotely located from the users because only the results of the execution are transmitted over the network to reduce bandwidth requirements and transmit time. As an example, several engineering and design teams work together on the design of a new vehicle. The design and simulation data are stored at a central storage site. At a design review, reviewers are

located at a number of locations remotely from the central storage site. The reviewers may review several different system designs of the vehicle. For example, one review team may review the exterior design of the vehicle, another may review a wind tunnel simulation on the vehicle exterior, yet another team may review the engine design. Centralized scalable resource system 10 is operable to access the different sets of design data stored at the central storage site, render graphical images of the design and then send the rendered graphical images to the respective review teams. The rendered images are displayed for viewing by the respective teams. The switching fabrics are operable to connect and route the output from the subset of compute resource units to the subset of visualization resource units that have been allocated to the job. Therefore, system 10 is scalable and reallocatable according to the requirements of the executed jobs.

卷之三